Available online at www.sciencedirect.com
ScienceDirect
Journal of Biomedical Informatics 39 (2006) 706–719
Journal of Biomedical Informatics
www.elsevier.com/locate/yjbin

A generic minimization random allocation and blinding system on web

Hongwei Cai [a], Jielai Xia [b,*], Dezhong Xu [c,*], Donghuai Gao [a], Yongping Yan [c]

[a] Network Center, Fourth Military Medical University, No. 17 Changle West Road, Xi'an, Shaanxi 710032, China
[b] Department of Health Statistics, Faculty of Preventative Medicine, Fourth Military Medical University, No. 17 Changle West Road, Xi'an, Shaanxi 710032, China
[c] Department of Epidemiology, Faculty of Preventative Medicine, Fourth Military Medical University, No. 17 Changle West Road, Xi'an, Shaanxi 710032, China

Received 1 November 2005
Available online 28 February 2006

Abstract

Background. Minimization is a dynamic randomization method for clinical trials. Although recommended by many researchers, the utilization of minimization has been seldom reported in randomized trials mainly because of the controversy surrounding the validity of conventional analyses and its complexity in implementation. However, both the statistical and clinical validity of minimization were demonstrated in recent studies. Minimization random allocation system integrated with blinding function that could facilitate the implementation of this method in general clinical trials has not been reported.

System overview. The system is a web-based random allocation system using Pocock and Simon minimization method. It also supports multiple treatment arms within a trial, multiple simultaneous trials, and blinding without further programming.

Methods. This system was constructed with generic database schema design method, Pocock and Simon minimization method and blinding method. It was coded with Microsoft Visual Basic and Active Server Pages (ASP) programming languages. And all dataset were managed with a Microsoft SQL Server database. Some critical programming codes were also provided.

Simulations and results. Two clinical trials were simulated simultaneously to test the system's applicability. Not only balanced groups but also blinded allocation results were achieved in both trials.

Discussions and conclusions. Practical considerations for minimization method, the benefits, general applicability and drawbacks of the technique implemented in this system are discussed. Promising features of the proposed system are also summarized.

Keywords: Random allocation; Minimization; Multi-treatment groups; WWW; Blinding

1. Introduction

The randomized controlled trials (RCTs) are commonly accepted as the gold standard research method for evaluating health care interventions. Randomization, which gives each subject a known (usually equal) chance of being assigned to any of the treatment groups, is the fundamental of RCTs [1]. Integrated with blinding, randomization helps to avoid possible bias in the selection and allocation of subjects arising from the predictability of treatment assignments [2]. The benefits of randomization can be summarized as "elimination of selection bias between groups, assurance of blinding and justification of randomization based tests" [1]. Although unrestricted randomization provides the best unpredictability and prevention of bias, it may yield highly disparate sample sizes or great imbalance on baseline characteristics between treatment groups by chance especially in smaller sample clinical trials [3,4]. Thus block randomization is used to ensure that comparison groups will be of approximately the same size. Stratified randomization, where a separate blocked randomization is performed in each stratum of interest, is one method of avoiding chance imbalances [5].

* Corresponding authors. Fax: +86 29 84774861 (J. Xia); +86 29 84774955 (D. Xu).
E-mail addresses: xiajielai@fmmu.edu.cn (J. Xia), xudezh@fmmu.edu.cn (D. Xu).

1532-0464/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2006.02.015

Minimization can be classified as a dynamic randomization method, as the allocation of the next subject is influenced by the current balance of the treatment groups. The minimization method was first proposed by Taves [6] in 1974 and independently described and generalized by Pocock and Simon [7] in 1975. The features of minimization methods and their potential problems have been comprehensively reviewed by Scott et al. [8] and McEntegart [9]. The advantages of minimization include "the balanced groups achieved with respect to both the numbers in each treatment arm and the characteristics of each group" and "the ability to incorporate more prognostic factors than for stratified randomization, which is particularly valuable in smaller sample trials" [8]. The drawbacks of using minimization include "controversy surrounding the validity of conventional analyses following minimization" [10,11], "selection bias due to the fact that next assignment can be predicted in some situations," and " the additional organizational complexity with the potential to harm recruitment and increase costs" [8]. However, these disadvantages are also true of other restricted allocation methods, and should not be weighted unduly [8].

The statistical comparison of minimization methods, stratified randomization and simple randomization was conducted by Hagino et al. [12] through simulations. Both the statistical and clinical validity of minimization were demonstrated in their studies. Although minimization is a highly effective method for treatment allocation and recommended by many commentators [12–14], the utilization of minimization was seldom reported [8].

Some random allocation software package or system using minimization methods had been developed, such as "Minim" [15], a free DOS (Disk Operating System, a command line user interface operating system) program for randomizing patients to treatment groups by the method of minimization, where data are saved in a text file. Multiple-user and multiple-center are not supported in this software, and the interface is not friendly. A web-based minimization random allocation system, which was developed by Kenjo et al. [16], is the most integrated and user friendly system available. It was developed by the Practical Extraction and Report Language (PERL) for writing common gateway interface (CGI) script. Multiple trials were not supported simultaneously, nor did it address the blinding of allocation results from participants. We have been engaged in the research of the minimization method. A minimization system on web based on conventional database structure for two treatments has been published [17]. The number of patients in each stratum and overall were well balanced between treatment groups in simulations, but only one clinical trial was supported at one time and the database structure should be modified if it was to be used in another clinical trial. Besides, it did not support blinding.

2. System overview

We have developed a random allocation and blinding system based on the Pocock and Simon minimization method that can be used in multiple clinical trials simultaneously. This system also supports multiple treatment arms within a trial and blinding without further programming.

2.1. System architecture

The system is running on two separate servers of Windows 2000 Server operating system. One is the World Wide Web (WWW) Server running Internet Information Services (IIS). The other is the Database Server running SQL Server 2000 which can not be accessed from Internet. The system is available, on-line, to WWW users through Secure Socket Layer (SSL), which provides the point-to-point communication security. The schematic of this system is shown in Fig. 1.

2.2. Definition of the relationships between users, centers, and studies

Two types of users, i.e., administrators and common users, have different privileges. Administrative tasks include viewing, editing, and deleting centers, users and designing studies. Common tasks include registering new patients and retrieving masked drug numbers.

A study is assumed to be either a single center study or a multi-center study. A user can only be allowed to login the system at one center, while one center may accommodate more than one user. A common user can participate in one or more studies from the center he resides. Patients

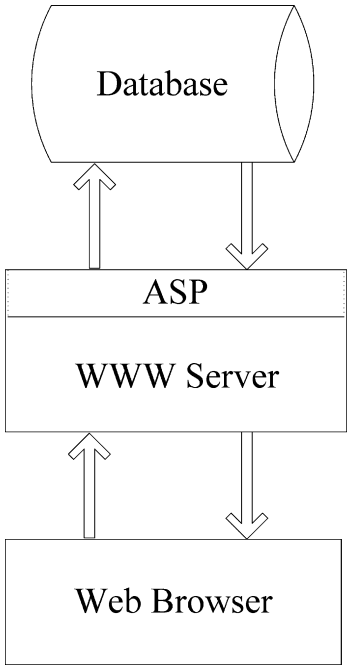


Fig. 1. Schematic of system architecture.

are registered into studies from certain center by common users. The tables demonstrating the relationships between users, centers and studies are shown in Fig. 2.

2.3. The management of this system

Managed by administrator, a new study can be initialized in the following approaches:

(1) Centers that will involve in the new study are added.
(2) Users in each center who will access this system as common users are produced.
(3) Name and parameters of the study are set, such as block length, treatment numbers, the bias assignment probability and common users who can register patients in this study through this system. For the sake of blind setting, block length should be integral multiple of treatment numbers so that there would be equal number of different drugs in one block. The important prognostic factors for the study and associated levels in each factor are identified and added first. Then Inclusion/Exclusion criteria of this study can be added with a serial number at the head of each item. The Inclusion/Exclusion criteria will be listed in web browser in ascending order of the serial numbers.
(4) Blinding. The number of patients that will be recruited in each participant center should be estimated first. Then consecutive masked drug numbers are produced according to the expected number of patients in each center. During the process of trials, new masked drug numbers would be added should the drug numbers in certain center be run out.

The management of this system is illustrated in Fig. 3.

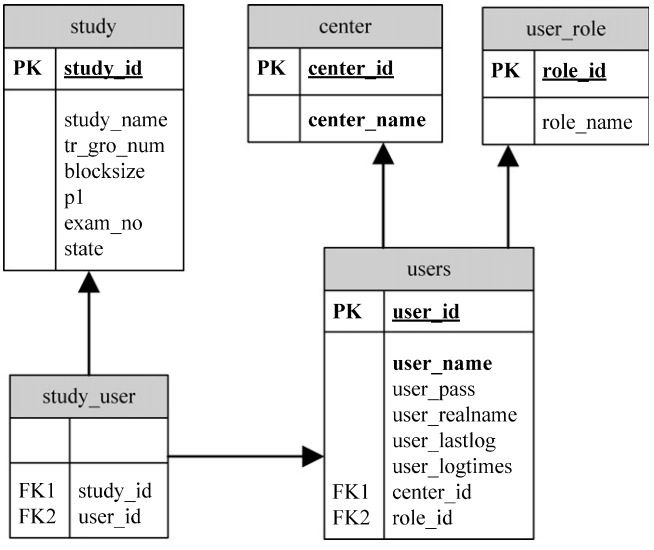


Fig. 2. Database structure representing relationships between users, centers and studies.

2.4. Patient recruitment and drug allocation process

All patients are recruited by common users. They use this system in the following steps after logging on

(1) Choose study and input simple demographic information about the new patient.
(2) Select items according to the Inclusion/Exclusion criteria and the system will automatically decide whether this patient is eligible or not to current study.
(3) If the patient is eligible, the levels for each prognostic factors associated with current patient are elected and submitted. Otherwise, the patient would be kept out of the randomization process.
(4) After the confirmation of submissions, a masked drug number is given back to this patient.

3. Methods

3.1. Database design

The database structure in this system makes it easy to accommodate new data items, such as trials, factors, and levels, without the additional programming that would be required in a traditional database design.

3.1.1. Relationships of clinical studies and patients

There may be several prognostic factors in one study and several levels in each factor. Each patient matches with certain level in any factor. Fig. 4 demonstrates how a female patient, aged 45 and enrolled from Southern hospital, is related with a study with 3 factors and 8 levels.

Little modification of the database is desirable when new clinical trial is added. A database structure improved from Entity-Attribute-Value (EAV) schema was used to show the relationship of clinical trials and subjects [18,19]. As illustrated in Fig. 5, each study might contain several prognostic factors, which were related with the study by Foreign Key of *study_id* in *factors* table. Each factor might contain several levels, which were related with the factor by Foreign Key of *level_id* in *levels* table. A row in *eav_pa_int* table represented a relationship between certain subject and level. With the matching of *patient_id* and *level_id*, every patient could be located with the exact levels. The patients having been allocated to certain treatment at certain level could be identified by column of "*treatment*" and "*level_id*" combination in *patient_en* table and *eav_pa_int* table. The number of patients at each level could be figured out by summing them up.

3.1.2. Study and Inclusion/Exclusion criteria

Minimization techniques utilize an adaptive schema, in which the next subject assignment is based on the current balance of the treatment groups. The appropriate drug number and associated treatment for the next

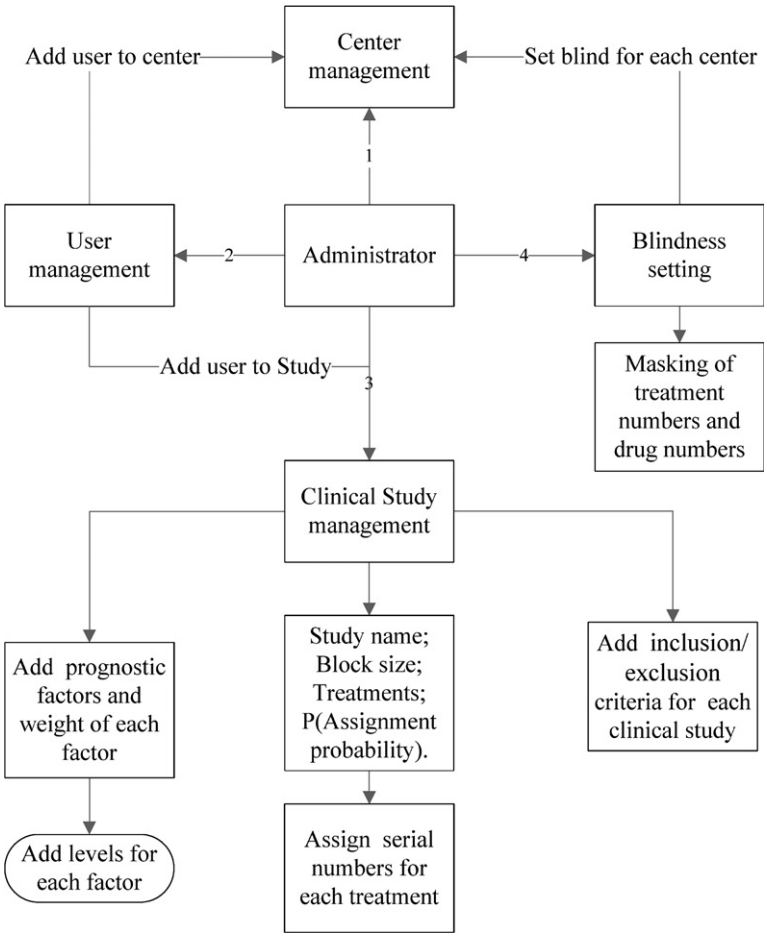


Fig. 3. Management of minimization random allocation system.

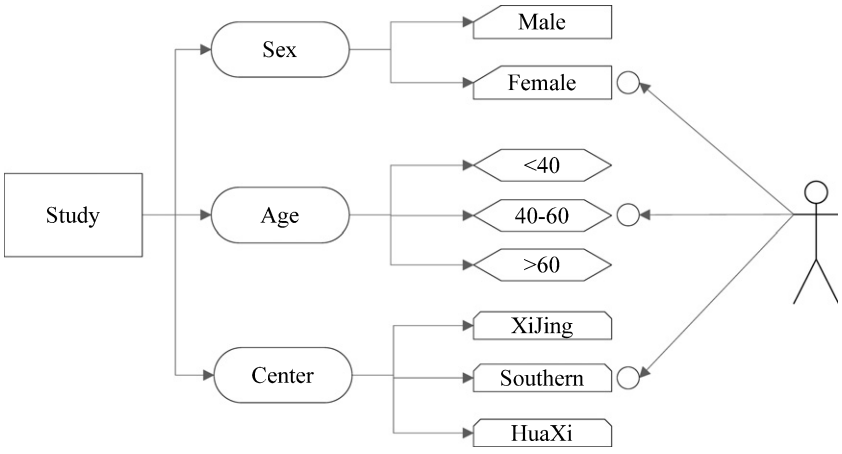


Fig. 4. Relationships between a study and a patient.

patient should only be allocated when the eligibility of the patient has been confirmed [20]. The database design showing the relationship between study and Inclusion/Exclusion criteria was referenced the EAV structure from TrialDB (TrialDB is an open-source Clinical Study Data Management System) [19]. This database structure, illustrated in Fig. 6, also supports multi-trial.

Arbitrary numbers of study criteria for certain study could be recorded in *study_criteria* table, where Inclusion and Exclusion criteria could be distinguished by *criterion_des*. Criteria for each study could be distinguished by *study_id* in *study_criteria* table. The responses (“0” stands for “no” and “1” stands for “yes”) of Inclusion/Exclusion criteria for all recruited patients were recorded in *patient_study_criteria* table.

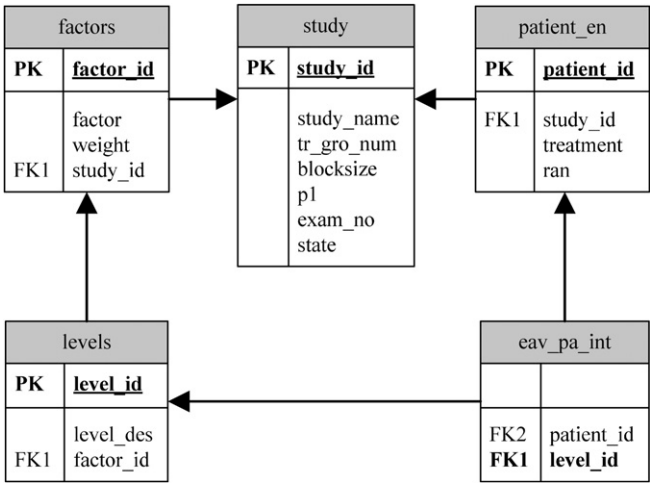


Fig. 5. Database structure representing relationships between enrolled patients and studies.

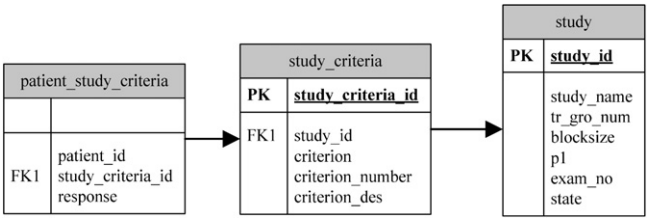


Fig. 6. Database structure representing the relationships between studies and Include/Exclude criteria.

3.2. The statistical method implemented in the computer program

The system was coded with Microsoft Visual Basic and Active Server Pages (ASP) programming languages, and the data were stored in a Microsoft SQL Server database. Programming was based on Pocock and Simon method, where treatment assignment depends on three aspects [7,8]: ① the amount of variation among assignments for any given factor level ② a measure of the total imbalance in treatment numbers ③ assignment probabilities to the arms of the trial.

Program flow chart of the computation and allocation process is shown in Fig. 7.

3.2.1. The computation and comparison of the amount of variation

Suppose patients are to be assigned to a clinical trial with N treatment groups and M prognostic factors for which treatment balance is required, the number of levels of these factors being $n_1, n_2, \ldots, n_M$. For an arbitrary point, let x_{ijk} be the number of patients with level j of factor i who have been assigned treatment k, $i=1, 2, \ldots, M$; $j = 1, 2, \ldots, n_i$ and $k = 1, 2, \ldots, N$. Suppose the newly registered and eligible patient is in level $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_M$ for factors $1, 2, \ldots, M$.

Define $\boldsymbol{G}_k$ as the amount of variation the patient brings suppose it to be assigned to certain treatment. If treatment k is assigned to the next patient, define $\boldsymbol{d}_{ik}$ as the resultant of "lack of balance" among treatment assignments for patients with level $\boldsymbol{r}_i$ of factor i. One-way to measure the total imbalance is to sum up the marginal imbalances. There may however be situations where some prognostic factors are considered to be more important than others. In such cases it is possible to use a weighted sum instead. The best weights for a given trial can be explored using simulations [9]. Then the amount of variation

$$\boldsymbol{G}_k = \sum_{i=1}^{M} w_{i*} d_{ik} \quad (w_i \text{ is the weighting of factor } i). \tag{1}$$

Define $\boldsymbol{S}_{ik}$, the *Standard Deviation* of the number of patients in each factor level, measures the "amount of variation" at level j in factor i.

$$\boldsymbol{d}_{ik} = S_{ik} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(x_{ik} - \overline{\mathrm{X}}_i)^2} \quad (i = 1, 2, \ldots, M.$$
$$k = 1, 2, \ldots, N.). \tag{2}$$

With the product of weighting and $\boldsymbol{S}_{ik}$ for each factor, Eq. (1) that measures the amount of variation becomes

$$\boldsymbol{G}_k = \sum_{i=1}^{M} w_{i*} S_{ik} = \sum_{i=1}^{M} w_{i*} \sqrt{\frac{1}{N}\sum_{k=1}^{N}(x_{ik} - \overline{\mathrm{X}}_i)^2}$$
$$(i = 1, 2, \ldots, M. \quad k = 1, 2, \ldots, N.). \tag{3}$$

The programming code to fulfill this function is given in Appendix A.

3.2.2. The choice of assignments probability and allocation

It was recommended that "Deterministic dynamic allocation procedures should be avoided and an appropriate effort should be made to be incorporated for each treatment allocation" [20]. Therefore, an appropriate assignment probability should be considered to the treatment arms of the trial.

In our system, the treatment assignment probability is unvaryingly equal to p if there is only one smallest $\boldsymbol{G}_k$. Suppose the bias treatment assignment probability is 0.8. As shown in Fig. 7, if there is only one smallest $\boldsymbol{G}_k$ and the random number generated by computer is less than 0.8, the target treatment is the one matched with the smallest $\boldsymbol{G}_k$. If the random number is equal to or more than 0.8, the treatment number matched with the smallest $\boldsymbol{G}_k$ would be out. Then assignments would proceed in other residual treatments in the same way. In other words, there is always a probability of 80% with which the treatment with the smallest $\boldsymbol{G}_k$ in the residual treatment number pool is chosen in each inning. If there is more than one treatment with the smallest and equal $\boldsymbol{G}_k$, the new patient would be assigned into these treatments with equal probability. If the assignment probability is set as 0.5 for a two-treatment clinical trial, patients will be allocated to treatments by unrestricted randomization method.

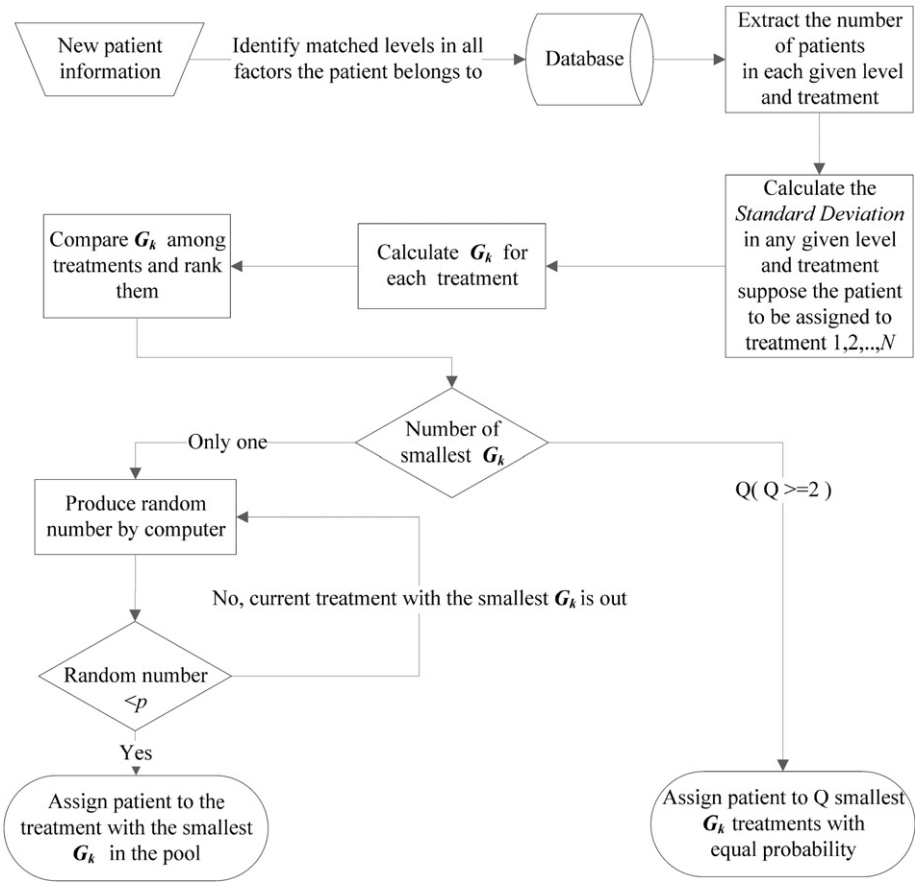


Fig. 7. Program flow chart of the computation and allocation process. G_k: the amount of variation, refer to Eq. (3) Random number: generated by computer and ranges from 0 to 1. p: the bias assignment probability.

The programming code for patient's allocation, written in Visual Basic Script and embedded in ASP, is given in Appendix B.

3.3. Blinding methods

The blinding information for all studies is stored in two tables. One is used for the matching of drug names and treatment numbers, where a kind of drug is assigned with a certain treatment number. The other is used for the matching of drug numbers and treatment numbers, where each drug number is associated with a treatment number randomly. The random matching method between drug numbers and treatment numbers, together with the drug allocation rule, are explained below.

3.3.1. Drug number producing for each treatment

A Stored Procedure in SQL Server 2000 database was written for the blinding of drug numbers and treatment numbers.

It takes four steps to fulfill this task in a block of drugs: ① *study_id* and *center_id* are read from web pages. The biggest *drug number*, *block number*, and *block size* are read from database. ② *Drug numbers* (for each trial, the value of drug number starts from 1 and with the increment of 1 automatically for the next drug), *block number* and *computer-generated random numbers* are inserted into the blind table. ③ *Drug allocation numbers* are updated in blind table according to the rank of random numbers from small to big ones. ④ *Treatment numbers* are allocated for each *drug number* according to the rank of random numbers from small to big ones. The number of different drugs in one block length is set equal.

For a clinical study with 3 treatment groups, if 6 is set as the block length, there would be 2 (6/3) drug numbers allocated for each treatment in one block. The blinding information for 7th to 9th block is demonstrated in Table 1. In the 7th block (from the 1st to 6th row in Table 1), the drug numbers (the column of "*Drug number*") are in ascending sort from 37 to 42. The rank of drug allocation numbers (the column of "*All_num*") are created according to the order of computer generated random numbers (the column of "*Ran*"). The smallest two drug allocation numbers are assigned to treatment 1, the middle two numbers are assigned to treatment 2, and the biggest two are assigned to treatment 3. In the end, there are two drugs for each treatment in the 7th block, and the matching between drug numbers and drug allocation numbers are randomized. Drug allocation numbers and associated treatment numbers are assigned to other drugs in the same way.

The programming code of the PROCEDURE for setting blind is given in Appendix C.

Table 1
Blind information for the 3rd center in a clinical trial

blind_id	center_id	study_id	patient_id	Drug number	Block number	Allocation number	Treatment number	Random number
1107	3	1	0	37	7	39	2	0.725715
1108	3	1	0	38	7	40	2	0.810006
1109	3	1	0	39	7	38	1	0.311874
1110	3	1	0	40	7	42	3	0.943732
1111	3	1	0	41	7	41	3	0.872126
1112	3	1	0	42	7	37	1	0.233886
1113	3	1	0	43	8	46	2	0.467334
1114	3	1	0	44	8	47	3	0.699734
1115	3	1	0	45	8	48	3	0.821765
1116	3	1	0	46	8	43	1	4.61E−02
1117	3	1	0	47	8	45	2	0.18797
1118	3	1	0	48	8	44	1	8.94E−02
1119	3	1	0	49	9	54	3	0.800063
1120	3	1	0	50	9	53	3	0.777643
1121	3	1	0	51	9	50	1	0.332821
1122	3	1	0	52	9	49	1	0.119092
1123	3	1	0	53	9	52	2	0.723484
1124	3	1	0	54	9	51	2	0.472167

3.3.2. Drug allocation to patients

Drugs tagged with drug numbers (each drug number is associated with a masked treatment) should be transferred to each participant center according to the *blind* table. They are allocated by the order of drug allocation numbers from small to big.

For example, the blinding information for a study is demonstrated in Table 1. Suppose drugs allocated to the 3rd center, where *center_id* is 3, are numbered from 37 to 54. The next patient registered from this center is to be allocated to treatment 1 by calculation of minimization method. In the blind table, 37 is the smallest drug allocation number associated with treatment 1 in this center in unallocated drugs where patient_id is "0" as being set as default. The drug number associated with that allocation number is 42. Therefore, drug number 42 is going to be allocated to this patient. Then the *patient_id* column in the corresponding row is updated with the new patient's *patient_id* to indicate that this drug has been allocated. The residual drug numbers are allocated to the following patients in the same way.

4. Simulations and results

4.1. Preparations before patients' recruitment

Two clinical trials were simulated in our system simultaneously.

There were two treatment groups and three prognostic factors known as "center, severity of disease and cardiopathy" in the first trial. "Severity of disease" was considered more important and the weighting for this factor was set two while the other two factors' weightings were set 1. There were 5 levels (center 1 to center 5) in factor "center," 3 levels (Light, Mild, and Severe) in factor "severity of disease" and 2 levels (Cardiopathy, No cardiopathy) in factor "cardiopathy." The bias treatment assignment probability p was set 0.8. Simulation data of 120 patients' were created by computer randomly. The total number of patients from center 1 to center 5 was 30, 18, 24, 22, and 26, respectively. As stated in Section 2.3, five centers were added as the participant centers in this study, adequate masked drug numbers were produced and allocated to each center in advance. Then five common users representing five centers were created, who would register patients from five different computers through web.

There were three treatment groups and three prognostic factors known as "sex, age, and center" in the second trial. The weightings were all the same for three prognostic factors. There were two levels (Male, Female) in factor "sex," three levels (<40, 40–60, and >60) in factor "age" and three levels (Xi Jing Hospital, Southern Hospital, and Hua Xi Hospital) in factor "center." The bias treatment assignment probability p was set 0.85. Simulation data of 54 patients' were created by computer. Three users representing three centers enrolled patients form three different computers through web.

4.2. Patients' recruitment and drug allocation

As stated in Section 2.4, common user logged on this system after its username and password being verified. Then trials' list appeared in which he had privilege to enroll patient. After a trial was chosen, new patient's demographic data would be input first. If the new patient should not meet the Inclusion/Exclusion criteria, a decline hint would appear. Otherwise, the randomization process began. After the patient's levels in all prognostic factors having been chosen and submitted, the treatment number that should be allocated to this patient was calculated by Pocock and Simon minimization method. Then according to the blinding rule, a masked drug number associated with the calculated treatment number would be given back to this patient.

In the first trial, for an arbitrary point in the assignment process, when the computer generated random number was less than $p(0.8)$, the treatment associated with the smallest G_k was allocated to the patient, the program output information of computing process is shown in Fig. 8. For an arbitrary point in the second trial, the program output information of computing process is shown in Fig. 9. The program output information is invisible to common users except the last line.

A common user's interface of drug allocation results through Internet Explore is shown in Fig. 10. He would have no idea of the real treatment applied to each recruited patient, for drug numbers and treatment numbers were randomly matched.

4.3. Allocation results of simulations

Both trials achieved the balanced groups with respect to both the numbers in each treatment arm and the characteristics of each group without mutual interference. The assignment result of the first simulated trial with 120 patients is shown in Fig. 11. The assignment result of the second simulated trial with 54 patients is shown in Fig. 12.

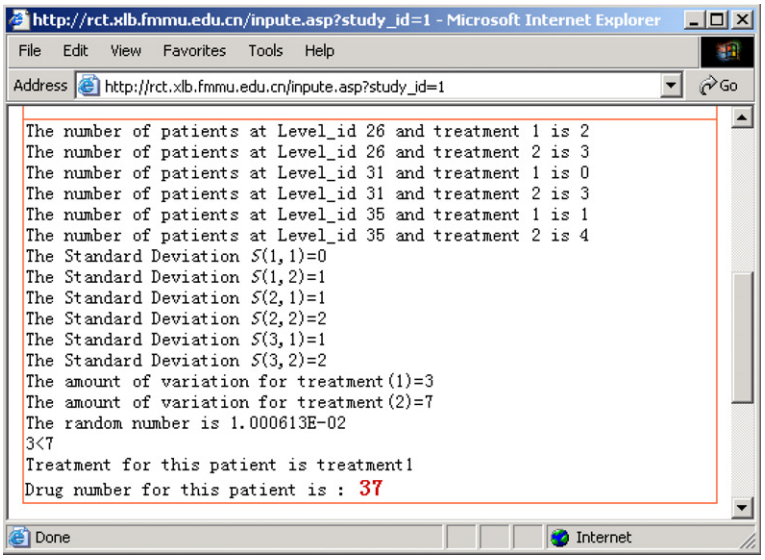


Fig. 8. The program output of the computing process and drug allocation result for the first trial.

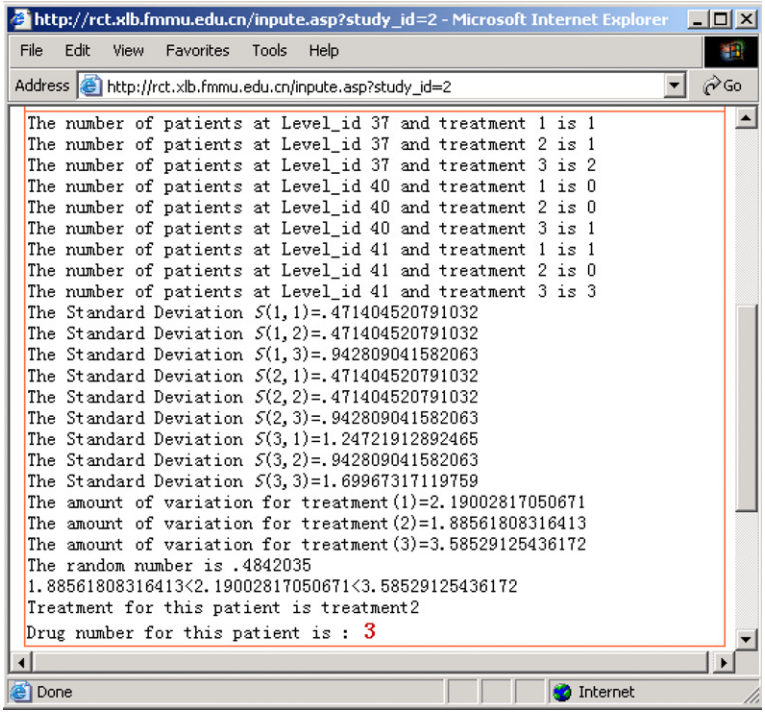


Fig. 9. The program output of the computing process and drug allocation result for the second trial.

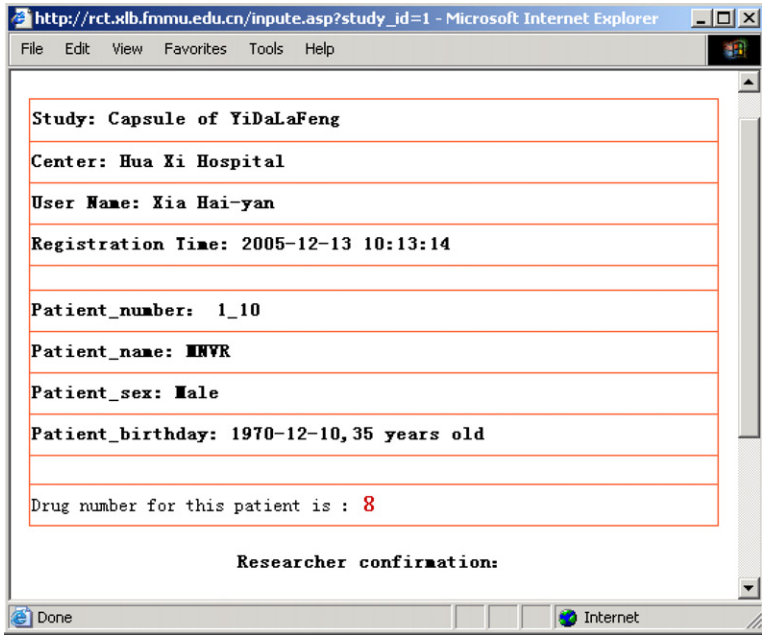
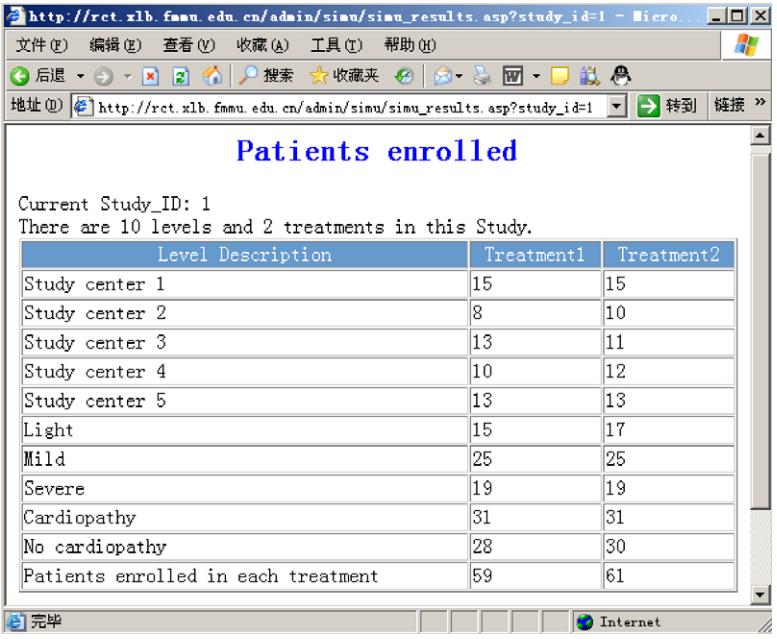


Fig. 10. Masked drug number is given back to the common user.



Level Description	Treatment1	Treatment2
Study center 1	15	15
Study center 2	8	10
Study center 3	13	11
Study center 4	10	12
Study center 5	13	13
Light	15	17
Mild	25	25
Severe	19	19
Cardiopathy	31	31
No cardiopathy	28	30
Patients enrolled in each treatment	59	61

Fig. 11. Allocation results of the first simulated trial with 120 patients.

5. Discussions

As a good car must be driven by a good driver to exert all its potential, the good manipulation of this system depends on the investigator's understanding of the minimization method. Practical considerations for minimization include:

(1) *Choice of random element.* The assignment probability is incorporated to reduce the predictability of next subject while keep good balance between treatment groups. However, the predictability of next subject is decreased with the decrease of the assignment probability while the imbalance probability among treatment arms increased at the same time. The proper assignment probability for a clinical trial could be varied depending on the trial size, number of treatments and number of stratifying factors [9]. Hagino et al. [12] simulated a team of 4-factor, 12-level and 2-treatment trials with varying assignment probabilities and sample sizes. The allocation probability of minimization was changed from 1.00 (deterministic

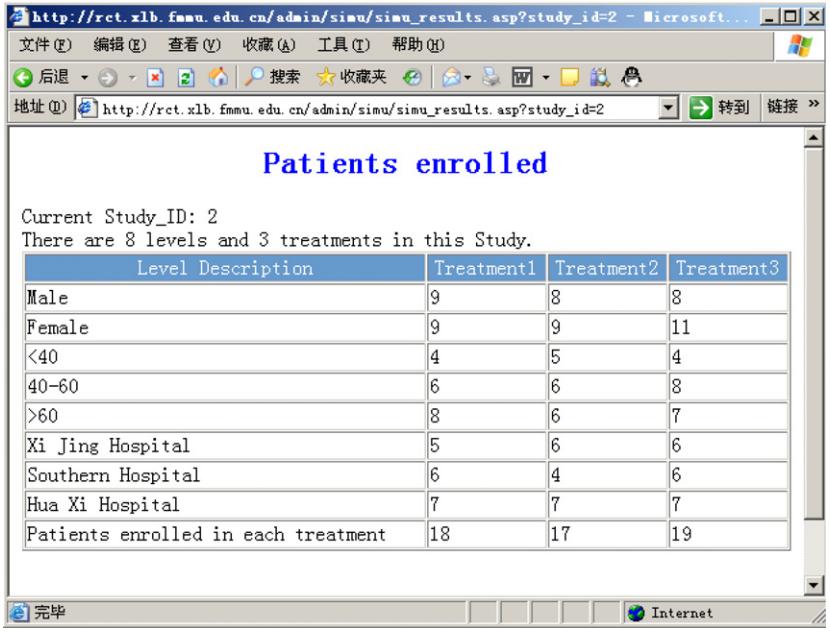


Fig. 12. Allocation results of the second simulated trial with 54 patients.

allocation) to 0.70 by 0.05 for 50, 100, 150, and 200 patients trails. Each combination was simulated 1000 times, respectively. The p value of theχ^2 test of the contingency table formed by simultaneous distribution of multiple prognostic factors and group had been calculated to evaluate balance in the simultaneous distribution. The bigger the p value of χ^2 test, the better the balance achieved. If the 1 percentile of p value of χ^2 test about 0.50 was chosen as a criterion to achieve a strongly acceptable degree of balance, the allocation probability of minimization was required to be set to 0.95 for 50, 0.80 for 100, 0.75 for 150, and 0.70 for 200 patients [12]. Proper allocation probability should be explored by simulations after the trial size, number of treatments, number of factors and levels are decided.

(2) *Choice of weightings.* Reed and Wickham concluded that there was no need to use weightings [21]. However, McEntegart argued that they could be useful to satisfy preferences about the relative merits of achieving balance at study [9]. It was advised that the best weightings for a trial could be explored using simulations [9].

Generic database structure in this system brings the benefit of keeping the database structure stable while accommodating arbitrary number of clinical trials with any number of treatments. The problem of representing phenotypic data is very similar to the problem of representing clinical patient data in this system [22,23]. The same modeling approach can be used in similar situations in other fields, such as genetics or genomics data management. One phenotype does not typically rely on a single gene, but rather the clustering of several genes and associate with the disparity in expression levels. A vast number of relative independent parameters (genotypes) can potentially associated with one or more patients (phenotypes). With the database design method in this system, the database structure could keep stable when new parameters (genotypes) or patients (phenotypes) being added. However, the database structure has its drawbacks. All dataset in *eav_pa_int* table needs to be scanned and accumulated time after time at the calculation of each new patient's assignment. Thus, response time of this system is mainly spent on the searches in *eav_pa_int* table. When the rows in this table increasing tremendously with new trials and patients added, the system response time may be a little postponed. Therefore, cleaning up of the trials' data is recommended after they have been accomplished. Another so called drawback about the database structure is that the patients' allocation data for a certain study must first be extracted into regular table before analyzed.

6. Conclusions

Promising features of the proposed system include: (1) New clinical trial can be added into this system without increasing the number of tables in database. (2) It supports arbitrary number of treatments in each trial. (3) Randomization allocation results can be masked without further programming. (4) Randomization phase begins only after the determination of subject eligibility for current trial. (5) Multi-center and multi-user are supported. (6) Participants can access the Internet-based automated system through web browser 24 h a day and 7 days a week.

This system described in this paper can be customized easily and facilitate the utilization of minimization method in the practice of clinical trials.

Appendix A. Programming code for computation and comparison of the amount of variation among treatments

```
Dim n 'The number of factors in current study
Dim p 'The number of treatment arms in current study
Dim w() 'Redim w(n) The weighting of each factors in current study
Dim s() 'Redim s(n,p) The number of patient in each factor levels and treatments already being recruited
Dim x() 'Redim x(n) Total patients' numbers at each factor levels
Dim ave() 'Redim ave(n) Average of patients' numbers at each factor levels
Dim sq() 'Redim sq(n,p) Sum of squares of "deviation from mean" at each factor levels
Dim sa() 'Redim sa(n,p) Standard deviation at each factor levels
Dim var() 'Redim var(p) Amount of overall imbalance suppose patient allocated to each treatment


// Calculate the standard deviation at each factor levels. (n,p is got from previous page, and the patient's number in each
factor levels and treatments already being recruited has been load into s(n,p).)


Redim x(n)
for i=1 to n
   x(i)=0
      for j=1 to p
         x(i)=x(i)+s(i,j)
      next
   ave(i)=(x(i)+1)/p

         for j= 1 to p
             sum=0
                   for h=1 to p
                       if h=j then
                       sum=sum+(s(i,h)+1-ave(i))*(s(i,h)+1-ave(i))
                       else
                       sum=sum+(s(i,h)-ave(i))*(s(i,h)-ave(i))
                       end if
                   next
              sa(i,j)=Sqr(sum/p)
         next
next


// Calculate the amount of overall imbalance suppose patient allocated to each treatment. (w(i) has been read from
database.)
for i=1 to p
    var(i)=0
      for j=1 to n
           var(i)=var(i)+sa(j,i)*w(j)
      next
next
```

Appendix B. Programming code for patients' allocation

```
Dim p1 'The assignment probability
Dim ran 'Random number created by computer
Dim var() 'Redim var(p) Amount of overall imbalance suppose patient allocated to each treatment
Dim a() 'Redim a(p) Record the initial position of each treatment when sorting the overall imbalance


for i =1 to p
      a(i)=i
```

```
next
call sort()

//Sort var(i) in ascending order, and at the same time, a(i) record its initial position in the array.
function sort()
Dim i,j,temp,temp_a
   for i = 1 to p
        for j = i + 1 to p
             If var(i) > var(j) Then
                  temp = var(i)
                  temp_a=a(i)
                  var(i) = var(j)
                  a(i)=a(j)
                  var(j) = temp
                  a(j)=temp_a
             end if
        next
   next

Dim m 'The number of smallest and equal G_k
Dim k 'The treatment group number to which the patient is to be allocated

m=1
for i=1 to p-1
    if var(i)=var(i+1) then
          m=m+1
    else
          exit for
    end if
next
if m=1 then
      for i=1 to p
          randomize() 'Use the return value of system timer function as the seed of randomization
          ran=rnd()
               If i=p or ran < p1 Then
                    k=a(i)
                    exit for
               End If
     next
else
   randomize()
   ran=rnd()
   r=int(ran*1000) mod m 'Get number from "0 to m-1" randomly
   k=a(r+1)
end if

//This patient would be allocated to treatment k
```

Appendix C. Programming code in SQL for setting blind

```
CREATE PROCEDURE preblind
@study_id smallint,
@center_id smallint
```

```
AS
DECLARE @medicine_num smallint – The biggest drug number for certain study and center.
DECLARE @block_num smallint – The biggest block number in current study.
DECLARE @counter smallint
DECLARE @t real – computer generated random number.
DECLARE @blocksize smallint – The block length in current study.
DECLARE @treatment smallint – The number of treatment arms in current study.
DECLARE @n smallint – The number of same treatment arms in each block.
DECLARE @m smallint

IF (select max(medicine_num) from blind where study_id=@study_id) is null
BEGIN
   SET @medicine_num=0
   SET @block_num=0
END
   ELSE
BEGIN
   select @medicine_num= (select max(medicine_num) from blind where study_id=@study_id)
   select @block_num= (select max(block_num) from blind where study_id=@study_id)
END
   select @blocksize = (select blocksize from study where study_id=@study_id)
   select @treatment = (select tr_gro_num from study where study_id=@study_id)
SET @counter=@medicine_num+1
WHILE @counter < =@medicine_num+@blocksize

– Insert random num into blind table
BEGIN

      select @t = RAND()
   BEGIN
   insert into blind (center_id,study_id,medicine_num,block_num,ran) values (@center_id,@study_id,@counter,
   @block_num+1,@t)
   END
   SET NOCOUNT ON
      SET @counter = @counter + 1
   SET NOCOUNT OFF
END
–insert allocation number into blind table according to the ranking of random number
SET @counter=@medicine_num+1
WHILE @counter < = @medicine_num +@blocksize
BEGIN
      IF @counter=@medicine_num+@blocksize
update blind SET allo_num=@counter where center_id=@center_id and study_id=@study_id and allo_num=0 and
medicine_num between (@medicine_num+1) and (@medicine_num+@blocksize)
      ELSE
   update blind set allo_num=@counter
where ran = (select min(ran) from blind where center_id=@center_id and study_id=@study_id and allo_num=0 and
medicine_num between (@medicine_num+1) and (@medicine_num+@blocksize))
   BEGIN
         SET NOCOUNT ON
         SET @counter = @counter + 1
         SET NOCOUNT OFF
   END
END
–insert treatment number into blind table according to the ranking of random number
SET @counter=@medicine_num+1
SET @n=@blocksize/@treatment
```

```
SET @m=1
WHILE @counter < = @medicine_num + @blocksize
BEGIN
    update blind set treatment=@m where allo_num=@counter
    BEGIN
        SET NOCOUNT ON
        SET @counter = @counter + 1
        SET NOCOUNT OFF
    END
    IF @counter > @medicine_num + @n*@m
    SET @m = @m + 1
END
```

References

[1] Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. Lancet 2002;359:515–9. Review.

[2] Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. Lancet 2002;359:696–700.

[3] Lachin JM. Properties of simple randomization in clinical trials. Control Clin Trials 1988;9:312–26.

[4] Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001;134:663–94. Review.

[5] Kernan WN, Viscoli CM, Makuch RW, Brass LM, Horwitz RI. Stratified randomization for clinical trials. J Clin Epidemiol 1999;52:19–26. Review.

[6] Taves DR. Minimization: a new method of assigning patients to treatment and control groups. Clin Pharmacol Ther 1974;15:443–53.

[7] Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 1975;31:103–15.

[8] Scott NW, McPherson GC, Ramsay CR, Campbell MK. The method of minimization for allocation to clinical trials. a review. Control Clin Trials 2002;23:662–74. Review.

[9] McEntegart DJ. The pursuit of balance using stratified and dynamic randomization techniques: an overview. Drug Inf J 2003;37:293–308.

[10] Halpern J, Brown Jr BW. Sequential treatment allocation procedures in clinical trials—with particular attention to the analysis of results for the biased coin design. Stat Med 1986;5:211–29.

[11] Lachin JM, Matts JP, Wei LJ. Randomization in clinical trials: conclusions and recommendations. Control Clin Trials 1988;9:365–74.

[12] Hagino A, Hamada C, Yoshimura I, Ohashi Y, Sakamoto J, Nakazato H. Statistical comparison of random allocation methods in cancer clinical trials. Control Clin Trials 2004;25:572–84.

[13] Blair E. Gold is not always good enough: the shortcomings of randomization when evaluating interventions in small heterogeneous samples. J Clin Epidemiol 2004;57:1219–22.

[14] Treasure T, MacRae KD. Minimisation: the platinum standard for trials? Randomisation doesn't guarantee similarity of groups; minimisation does. Br Med J 1998;317:362–3.

[15] <http://www.sgul.ac.uk/depts/phs/guide/minim.htm>; 2005 [accessed].

[16] Kenjo Y, Antoku Y, Akazawa K, Hanada E, Kinukawa N, Nose Y. An easily customized, random allocation system using the minimization method for multi-institutional clinical trials. Comput Methods Programs Biomed 2000;62(1):45–9.

[17] Cai HW, Xia JL, Xu DZ, Gao DH, Zhang Y. A random allocation system using minimization method on Internet. J Fourth Mil Med Univ 2004;25:1278–80. In Chinese.

[18] Brandt CA, Morse R, Matthews K, Sun K, Deshpande AM, Gadagkar R, et al. Metadata-driven creation of data marts from an EAV-modeled clinical research database. Int J Med Inform 2002;65:225–41.

[19] <http://ycmi.med.yale.edu/trialdb/>; 2005 [accessed].

[20] Lewis JA. Statistical principles for clinical trials (ICH E9): an introductory note on an international guideline. Stat Med 1999;18:1903–42.

[21] Reed JV, Wickham EA. Practical experience of minimization in clinical trials. Pharmaceut Med 1988;3:349–59.

[22] Negm RS, Verma M, Srivastava S. The promise of biomarkers in cancer screening and detection. Trends Mol Med 2002;8:288–93. Review.

[23] Baumgartner C, Matyas G, Steinmann B, Eberle M, Stein JI, Baumgartner D. A bioinformatics framework for genotype–phenotype correlation in humans with Marfan syndrome caused by FBN1 gene mutations. J Biomed Inform 2005 Jul 29; [Epub ahead of print].